

520.43032X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): SATOYAMA, et al.
Serial No.: Not assigned
Filed: August 28, 2003
Title: STORAGE SYSTEM
Group: Not assigned

LETTER CLAIMING RIGHT OF PRIORITY

Mail Stop Patent Application
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

August 28, 2003


Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Application No.(s) 2003-178879 filed June 24, 2003.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl J. Brundidge
Registration No. 29,621

CIB/amr
Attachment
(703) 312-6600

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 6月24日
Date of Application:

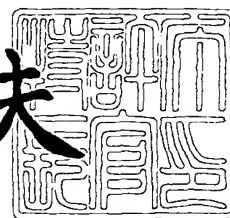
出願番号 特願2003-178879
Application Number:
[ST. 10/C]: [JP2003-178879]

出願人 株式会社日立製作所
Applicant(s):

2003年 8月 1日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



出証番号 出証特2003-3061648

【書類名】 特許願

【整理番号】 K02017711A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 12/00

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 里山 愛

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 山本 康友

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 江口 賢哲

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I D システム事業部内

【氏名】 大野 洋

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶装置システム

【特許請求の範囲】

【請求項 1】

複数の記憶装置と、前記複数の記憶装置を制御する記憶制御装置を備える記憶装置システムにおいて、

前記記憶制御装置は、計算機と接続するための第一のインタフェースと、プログラムを格納するメモリと、前記プログラムを実行する CPU と、前記複数の記憶装置に接続するための第二のインタフェースを備え、

前記プログラムは、前記記憶装置内のデータをコピーする際、コピー元ボリュームの特性に応じて、コピー先ボリュームを決定する部分を備える、記憶装置システム。

【請求項 2】

前記記憶装置システムは、複数種類の記憶装置を備え、

前記プログラムは、前記コピー元ボリュームを格納する記憶装置の種別に応じて、コピー先ボリュームを決定する部分を更に備える、請求項 1 記載の記憶装置システム。

【請求項 3】

前記記憶装置システムは、前記計算機からの入出力データを一時的に格納するキャッシュを備え、

前記プログラムは、コピー元ボリュームが前記キャッシュ内の記憶領域に割り当てられている場合、前記記憶領域の割当状況に応じて、コピー先ボリュームを決定する部分を更に備える、請求項 1 記載の記憶装置システム。

【請求項 4】

前記プログラムは、コピー元ボリュームの特性が複数ある場合、前記メモリに予め設定された特性の優先順位情報に従って、コピー先ボリュームを決定する部分を更に備える、請求項 1 記載の記憶装置システム。

【請求項 5】

前記記憶装置システムは、管理端末と接続するための第三のインタフェースを

備え、

前記プログラムは、コピー元ボリュームの特性が複数ある場合、前記管理端末からの指示に従い、コピー先ボリュームを選択する部分を更に備える、請求項 1 記載の記憶装置システム。

【請求項 6】

前記メモリは、コピー先ボリュームの選択基準と当該基準によるコピー先ボリュームの判定の仕方を記載した、基準情報を格納し、

前記プログラムは、前記基準情報に基づいて、コピー先ボリュームを決定する部分を更に備える、請求項 5 記載の記憶装置システム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、記憶装置システムに関し、特に、データの複製（以下「レプリケーション」）作成時、コピー先となるボリュームを選択可能な記憶装置システムに関する。

【0 0 0 2】

【従来の技術】

近年、企業が保有する記憶装置システム内の記憶装置に格納されたデータのレプリケーションを、別の記憶装置に作成する処理（以下「バックアップ」）に係る時間を短縮したいという要求が高まっている。これは、情報量の増加に伴いバックアップに係る時間が益々増加する一方で、企業の業務時間の延長により、バックアップの処理に割り当てられる時間が短縮していることが背景にある。そこで、現在、企業において、バックアップすべきデータのレプリケーションを、別の記憶領域に作成し、ユーザは元の記憶領域を業務で使用し、コピーされた記憶領域に格納されたデータをバックアップに使用する技術が用いられている。これにより、日常の業務を停止せずに、記憶装置システムに格納されたデータをバックアップすることができる。

【0 0 0 3】

記憶装置には、他の機器との接続に使用するインタフェースが異なるいくつか

の種類がある。例えば、ANSI T11によって標準化作業が行なわれているファイバチャネル (Fibre Channel、以下「FC」と称する) インタフェースを備えた記憶装置 (以下「FC記憶装置」と称する) やANSI T13によって標準化作業が行なわれているATA (AT Attachment) インタフェースを備えた記憶装置 (以下「ATA記憶装置」と称する) がある。

【0004】

ATA記憶装置は、比較的安価で、家庭用のデスクトップコンピュータなどに利用される。一方、FC記憶装置は、ATA記憶装置と比較して、データ入出力 (以下「I/O」) の処理能力が高く、24時間365日の連続動作に耐える高い信頼性を有するため、信頼性が要求される業務サーバなどで利用される。

【0005】

これらの記憶装置をディスクアレイ等の記憶装置システムに使用する場合、記憶装置システムに要求される性能、価格等に応じて記憶装置を使い分けることがある。例えば、記憶装置システムに格納されるデータの信頼性を向上する技術が、特許文献1に開示されている。また、高価で高信頼・高性能の記憶装置と安価で信頼性の低い記憶装置とを一つの記憶装置システムで混在させて用途毎に使い分ける技術と、レプリケーション作成時にコピー先ボリュームを決定する技術が、特許文献2に開示されている。

【0006】

更に、データの種類毎に区分化されたキャッシュ領域を持つ場合に、その領域割り当てを最適化してキャッシュヒット率を向上させる技術が、特許文献3に開示されている。

【0007】

【特許文献1】

特開平10-301720号公報

【特許文献2】

特願2003-37602号抄録

【特許文献3】

米国特許 第5434992号公報

【0008】**【発明が解決しようとする課題】**

記憶装置システムにおいてレプリケーションを作成する際、記憶装置システム内に種類の違う記憶装置を持ったり、キャッシュを分割したりといった特性がある場合、コピー元ボリュームに関するそれらの特性との相関を考慮してコピー先ボリュームの候補を決定することについて考慮されていない。また、記憶装置システムが大容量になる程、コピー先ボリュームの候補数が増大するので、コピー先ボリュームを決定する際のユーザの負担も増大するが、ユーザの負担を軽減することが考慮されていない。

【0009】

本発明の目的は、レプリケーションを作成する際、コピー元の特性を考慮してコピー先を決定する記憶装置システムを提供することにある。また、ユーザの負担を軽減するため、ユーザが前記特性を意識せずにコピー先を選択できるようなレプリケーション支援機能を有する記憶装置システムを提供することにある。

【0010】**【課題を解決するための手段】**

上記課題を解決するため、コピー元ボリュームが格納されている記憶装置の種類や、コピー元ボリュームのキャッシュの記憶領域の割当状況に応じて、コピー先ボリュームを決定する。また、コピー先ボリュームの選択基準と当該基準によるコピー先ボリュームの判定の仕方を記載した基準情報を用いて、コピー先ボリュームを決定する。

【0011】**【発明の実施の形態】**

図1は、コピー元ボリュームの特性を考慮してコピー先ボリュームを選択できる記憶装置システムを含む計算機システムの実施例を示す図である。図において、プログラムや情報は丸で囲んで示す。

【0012】

第一の記憶装置システム70Aは、ホスト10、管理サーバ100及び第二の記憶装置システム70Bと接続されている。ホスト10等との接続インターフェー

スとしては、FCやSCSI等が使用される。管理サーバ100は、管理用ネットワーク200を介して、第二の記憶装置システム70Bは、通信路300を介して接続される。尚、通信路300には、例えばFCやESCONなどが考えられるが、本実施形態では特に区別しない。

【0013】

第一の記憶装置システム70Aは、記憶制御装置20、FC記憶装置群28及びATA記憶装置群31とから構成される。

【0014】

記憶制御装置20は、CPU21、メモリ22、ホスト10からの入出力データを一時的に格納するキャッシュ23、ホスト10と記憶制御装置20との間のデータ送受信の制御を行うホストFC I/F24、通信路300との間のデータ送受信の制御を行う記憶装置システムFC I/F14、FC記憶装置群28との間のデータの送受信を制御する記憶装置FC I/F25、ATA記憶装置群31との間のデータの送受信を制御する記憶装置ATA I/F26、管理サーバ100との間の情報の送受信を制御する管理I/F27を有し、これらは、内部バスで相互に接続される。

【0015】

FC記憶装置群28は、複数のFC記憶装置29を、ATA記憶装置群30は、複数のATA記憶装置32を有する。尚、本発明では、記憶装置システムに採用される記憶装置はこれらの記憶装置に限られない。ATA記憶装置にはSATA (Sequential ATA) 記憶装置等がある。

【0016】

記憶装置FC I/F25は、FC記憶装置群28とFCを介して接続される。尚、FCに用いるプロトコルとして、FCアービトレーションループやFCのポイント・ツー・ポイント、ファブリック等の他のプロトコルでもよい。

【0017】

記憶装置ATA I/F26は、ATA記憶装置群31とATAバスを介して接続される。

【0018】

メモリ 22 には、CPU 21 で実行される種々のプログラムが格納される。具体的には、記憶装置システム 70 の動作を制御するための RAID 制御プログラム 41 及び記憶装置システム 70 の構成を管理するための管理エージェント 80 である。又、メモリ 22 には、各種管理情報も格納される。具体的には、FC 記憶装置群 28 と ATA 記憶装置群 31 に関する情報を記録する記憶装置管理テーブル 44、FC 記憶装置群 28 に作成される論理的記憶領域（以下「FC LU」）30 及び ATA 記憶装置群 31 に作成される論理的記憶領域（以下「ATA LU」）33 に関する情報を記録する LU 管理テーブル 45、データのコピー元とコピー先に関する情報を記録するペア管理テーブル 46、キャッシュグループ（後述）の管理を行うためのキャッシュグループ情報テーブル 47、記憶装置システム 70 B が記憶装置システム 70 A に対し自身の LU を記憶装置システム 70 A の LU として提供するための構成情報テーブル 48 を有する。

【0019】

尚、キャッシュグループとは、FC LU 30、ATA LU 33 において、LU をもとに記憶領域のグループ化が行われ、そのグループ毎にキャッシュ 23 内の記憶領域が割り当てられたときの、LU のグループのことを指す。

【0020】

RAID (Redundant Array of Inexpensive Disks) 制御プログラム 41 は、FC 記憶装置群 28 と ATA 記憶装置群 31 にコマンドを発行する部分（不図示）と、FC 記憶装置群 28 と ATA 記憶装置群 30 の管理を行う部分（不図示）と、各記憶装置群に設定された論理的記憶領域（以下「LU」）の管理を行う部分（不図示）を有する。RAID 制御プログラム 41 内には、サブプログラムとして、記憶装置システム 70 内でデータのレプリケーションを作成するためのレプリケーション作成プログラム 42 と、副 LU 選択支援プログラム 43 とがある。尚、データのレプリケーションを実施する際、同期（データのコピーの完了を待って、上位装置に完了を報告する）、非同期（データのコピーの完了を待たずに、上位装置に完了を報告する）といったバリエーションがあるが、本実施形態では、特に区別しない。

【0021】

管理エージェント 80 は、管理サーバ 100 からの入力を受けて、記憶装置に関する情報（以下「記憶装置情報」）を設定したり、記憶装置情報を管理端末 100 へ出力するためのプログラムである。

【0022】

管理サーバ 100 は、CPU 110、主記憶 120、入力部（キーボード装置等）130、出力部（ディスプレイ装置等）140、管理用ネットワーク 200 と接続するための管理 I/F 150 及び、記憶部 160 を有し、これらは、内部バスで相互に接続される。記憶部 160 には、CPU 110 で実行されるプログラムであるストレージマネージャ 131 及び副 LU 選択支援プログラム 43 が格納される。CPU 110 は、これらのプログラムを実行して、記憶装置システム 70 が有するテーブル（44～48）に格納された情報を定期的に収集し、これらのテーブルのレプリケーションを作成する。

【0023】

ホスト 10 は、例えば、パソコンやワークステーション、汎用コンピュータなどであり、外部に接続するための FC インタフェースである HBA（Host Bus Adaptor）（不図示）を備える。HBA にも WWN が付与されている。

【0024】

図 2 は、記憶装置管理テーブル 44 の一例を示す図である。本テーブルは、記憶装置毎に存在する情報であり、記憶装置番号 241、記憶装置種別 242、アレイ構成 243、用途 244、起動状況 245 がある。

【0025】

記憶装置番号 241 は、それぞれの記憶装置に割り振られた番号を示す。

【0026】

記憶装置種別 242 は、記憶装置の種別（FC、ATA 等）を示す。

【0027】

アレイ構成 243 は、記憶装置が含まれる RAID グループ（冗長性を構成する記憶装置の組）を作成した順番を示す番号と RAID レベルを示す。例えば、「（1）RAID 5」が登録された記憶装置は、1 番目に作成された RAID レベル 5 の RAID グループに含まれる記憶装置である。尚、記憶装置システム 7

0では複数のRAIDグループを作成すること、例えばRAID1とRAID5のRAIDグループを作成することが可能である。又、RAIDグループに含まれる記憶装置は、記憶装置システム70が有する記憶装置全てでも一部でもよい。

【0028】

用途244は、記憶装置が含まれるRAIDグループのユーザの用途（データベース用途を示すDB、ファイルシステム用途を示すFS等）を示す。

【0029】

起動状況245は、記憶装置が起動しているか否か（起動している場合はON、停止している場合はOFF）を示す。

【0030】

図3は、LU管理テーブル45の一例を示す図である。本テーブルは、記憶制御装置20が各LUを管理するために各LU毎に存在する情報であり、LU番号251、ホスト割当状況252、LUN253、容量254、LU種別255、相手LU番号256がある。

【0031】

LU番号251は、各LUに対して任意に割り当てられた識別子を示す。

【0032】

ホスト割当状況252は、LUがホスト10に割り当てられているか否か（割り当てられている場合は「有」、割り当てられていない場合は「無」）を示す。

【0033】

LUN253は、LUがホスト10に割り当てられている場合、ホスト10が当該LUをアクセスするために必要なSCSIロジカルユニット番号を示す。

【0034】

容量254は、LUに割り当てられている記憶容量を示す。

【0035】

LU種別255は、LUの種別（FC LU32かATA LU33）を示す。

【0036】

相手LU番号45は、オリジナルのデータが格納されたLU（以下「正LU」

）ならば当該データのコピーされたデータが格納されたLU（以下「副LU」）の、副LUならば正LUの、任意に割り当てられた識別子を示す。

【0037】

図4は、ペア管理テーブル46の一例を示す図である。本テーブルは、記憶装置システム70内の、コピーされたデータを保持するLUの対（以下「LUペア」）を管理するための情報であり、LUペア番号261、正LU情報262、副LU情報263、ペア状態264がある。

【0038】

LUペア番号261は、LUペアに対して任意に割り当てられた識別子を示す。

【0039】

正LU情報262は、識別子が与えられたLUペアのうち、正LUに割り振られたLU番号を示す。

【0040】

副LU情報263は、識別子が与えられたLUペアのうち、副LUに割り振られたLU番号を示す。

【0041】

ペア状態264は、LUペアの現在の状態を示す。例えば、LUペアの各々のLUに格納されているデータの同期が取られ、各々に格納されたデータの内容が一致している状態（以下「Pair状態」と称する）又は、LUペア間でデータの同期が取られていない状態（以下「Split状態」と称する）などがある。

【0042】

記憶装置システム70Aは、例えば、Pair状態にあったLUペアを、任意の時間においてSplit状態に変更する。このとき、副LUには、任意の時間におけるLUペアが有するデータが保存される（このような処理を「スナップショットを取る」という）。その後、ホスト10が副LUからデータを読み出して他の記憶装置（例えばテープ装置等）に書き込むことで、スナップショットを取った時間におけるLUペアに格納されたデータのバックアップをとることができる。尚、スナップショットを取った後の副LUそのものをデータのバックアップ

として保存しておいてもよい。

【 0 0 4 3 】

以下、記憶制御装置 2 0 が、ユーザ又は管理者の指示に従って、記憶装置管理テーブル 4 4 を用いて、R A I D グループの構築を行う場合について説明する。

【 0 0 4 4 】

記憶装置システム 7 0 に電源が投入された際に、C P U 2 1 は、R A I D 制御プログラム 4 1 を実行して、記憶装置 F C I / F 2 5 及び記憶装置 A T A I / F 2 6 に接続されている記憶装置を探索し、検出された記憶装置について、記憶装置管理テーブル 4 4 の記憶装置番号 2 4 1 及び記憶装置種別 2 4 2 に、該当する情報を登録する。

【 0 0 4 5 】

尚、記憶装置の種類は、ユーザが管理サーバ 1 0 0 の入力部 1 3 0 から入力してもよい。この場合、C P U 2 1 は、管理サーバ 1 0 0 から受け取った情報に基づいて、記憶装置管理テーブル 4 4 に情報を登録する。

【 0 0 4 6 】

次に、C P U 2 1 は、ユーザの指示に基づいて、アレイ構成 2 4 3 や用途 2 4 4 の設定を行う。まず、ユーザが入力部 1 3 0 から入力した記憶装置番号 2 4 1 、記憶装置種別 2 4 2 の情報取得の指示を受信すると、C P U 2 1 は R A I D 制御プログラム 4 1 を実行して、要求された情報を記憶装置管理テーブル 4 4 から取得し、管理 I / F 2 7 を介して管理サーバ 1 0 0 に転送する。管理サーバ 1 0 0 は、受信した情報を、出力部 1 4 0 に表示する。尚、ユーザが記憶装置の種別を定義する場合には、本処理は省略されてもよい。

【 0 0 4 7 】

その後ユーザは、出力部 1 4 0 に表示された情報に基づいて任意の記憶装置を選定し、選定された記憶装置を用いて R A I D グループを構築する指示を入力部 1 3 0 から入力する。また、ユーザは、構築する R A I D グループの用途を入力する。C P U 2 1 は、管理 I / F 2 7 を介して受信した R A I D グループや用途等の情報を、記憶装置管理テーブル 4 4 のアレイ構成 2 4 3 、用途 2 4 4 に登録する。

【 0 0 4 8 】

図 5 は、キャッシュグループ情報テーブル 4 7 の一例を示す図である。本テーブルは、記憶装置システム 7 0 A がキャッシュグループの管理に用いる情報で、キャッシュグループの識別子であるキャッシュグループ ID 4 6 1、キャッシュグループに対して割り当てられるキャッシュ量であるキャッシュ量 4 6 2、キャッシュグループに属する LU の識別子である LU ID 4 6 3 がある。本テーブルにより、キャッシュグループの作成・削除や、それに所属する LU の追加・削除は動的（「他の処理を停止させることなく実施」を意味する）に行うことができる。また、キャッシュグループへのキャッシュ 2 3 の割当量を動的に変更する機能を有する。

【 0 0 4 9 】

図 6 は、構成情報テーブル 4 8 の一例を示す図である。本テーブルには、記憶装置システム 7 0 A が扱う LU に関する情報が記述されており、LU が接続している外部インタフェースのポートの ID であるポート ID 4 8 1、ポート ID に対応する WWN 4 8 2、各 LU の LUN（Logical Unit Number）である LUN 4 8 3、記憶装置システム 7 0 A および 7 0 B の LU が提供する記憶領域の容量である容量 4 8 4、マッピング情報 4 8 5 がある。

【 0 0 5 0 】

マッピング LUN 4 8 5 は、そのポートおよび LUN に対応づけられている記憶装置システム 7 0 B の LU の識別子である。つまり、マッピング LUN 4 8 5 に記述のある場合は、その LU は記憶装置システム 7 0 B の LU であることを示し、それ以外の LU は、記憶装置システム 7 0 A の LU であることを示す。

【 0 0 5 1 】

記憶装置システム 7 0 A は、構成情報テーブル 4 8 を用い、記憶装置システム 7 0 B の LU を、あたかも記憶装置システム 7 0 A の LU であるかのように、ホスト 1 0 に提供する。つまり、ホスト 1 0 は、記憶装置システム 7 0 B の LU に対するデータ入出力要求を、記憶装置システム 7 0 A に対して行うことができる。

【 0 0 5 2 】

図8は、正LUに対して最適な副LUを自動的に（副LUを選択するにあたり、条件を指定しない）割り当てる処理フローである。

【0053】

CPU21は、副LUを選択可能な範囲において、候補となる空きLUが存在するか否かを判定する（ステップ801）。ここで選択可能な範囲とはレプリケーションを作成するホストに割り当てたLUとする。空きLUがない場合は、出力部52にメッセージを出力して（ステップ802）、終了する。空きLUがある場合は、正LUの特性を記憶装置システム70A内の記憶装置管理テーブル44及びLU管理テーブル45から検索する（ステップ803）。検索した正LUと同じ特性を持つLU群（以下「特性グループ」）に空きLUがある場合（ステップ804）、空きLUの中から1つをランダムに選択し、副LUとして割り当てる（ステップ805）。例えば、同じ記憶装置種別（FC記憶装置/ATA記憶装置）を同一特性グループとし、正LUが格納される記憶装置の記憶装置種別と同じ特性グループの中から空きLUを検索する。他の例として、同じキャッシュグループを割り当てられたLUを同一特性グループとし、正LUのキャッシュグループと同じグループのLUの中から空きLUを検索する。更に他の例として、同じ記憶制御装置20を同一特性グループとし、正LUを制御する記憶制御装置20配下のLUの中から空きLUを探す。また、正LUの特性グループが複数種類ある場合、正LUの複数の特性と同じ特性をもつものを同一特性グループとし、その中から空きLUを検索する。

【0054】

ステップ805の後、当該副LUについて、LU管理テーブル45及びペア管理テーブル46に登録する（ステップ807）。

【0055】

ステップ804で空きLUがないと判定されたら、他の特性グループから空きLUを選択する（ステップ806）。

【0056】

図9は、図8のステップ806の詳細手順を示す処理フローである。

【0057】

CPU 21は、正LUの特性の種類が複数あるか否か判定する（ステップ901）。特性の種類が1つの場合、正LUと違う特性グループから1つを選択する。例えば正LUがFC記憶装置であればATA記憶装置などのFC記憶装置以外から選択する。正LUの特性と違う特性グループが複数個ある場合の選択の基準は、空きLUがあり、かつ、負荷の軽いグループを選択する（ステップ902）。負荷の軽いか否かの判定基準は、その特性グループ内におけるレプリケーション作成数が少ないものを負荷が軽いと判断する。次に、CPU 21は、選択したグループ内の空きLUをピックアップし（ステップ904）、ピックアップした空きLU群の中から1つをランダムに選択する（ステップ905）。副LUを選択したら、当該副LUについて、LU管理テーブル45及びペア管理テーブル46に登録する（ステップ906）。ステップ901で正LUの特性の種類が複数個あると判断された場合、正LUが持つ複数の特性のうち、空きLUのある特性の一つを記憶装置システム70側でランダムに選択する（ステップ903）。例えば、記憶装置種別及びキャッシュ分割単位の特性を持つ正LUの場合、両方の特性を持つ空きLUがない場合、記憶装置種別が同じで別のキャッシュ分割単位のグループから空きLUを選択する。キャッシュ分割単位のグループが複数ある場合は、ステップ902同様、負荷の軽いグループを選択する。

【0058】

次に、副LUを選択するにあたり、ユーザが条件指定する場合について説明する。

【0059】

図10に特性の優先順位情報テーブル1300の一例を示す。本テーブルはメモリ22に格納されており、特性の優先順位はユーザに予め設定される。CPU 21は、当該優先順位を参照して、特性グループを決定する。従って前記例の正LUと記憶装置種別が同じでキャッシュ分割単位が異なるグループなのか、正LUと同じキャッシュ分割単位で記憶装置種別が違うグループなのか等を記憶装置システム70側で判定する。更に、ユーザが判定基準を入力指示してもよい。例えば、ユーザがコピー性能を重視したいことを入力指示すると、正LUと同一のキャッシュ単位から空きLUを選択する。また、副LUのコピー処理が正LUへの

I/O性能へ影響させることを防止するため、正LUの性能と副LUとの性能を独立させたい場合、正LUと異なるキャッシュ単位から空きLUを選択する。

【0060】

ユーザから条件の指定がない場合は、記憶装置システムが本テーブルの優先順位に従い副LUを選択する。この例では、優先順位が一番高い項目はハードディスク種別1301である。従って、副LUは正LUと同じ種類の記憶装置（FCかATA）を選択する。その中で次の優先順位に当てはまるものがあればそれを選択する。この例ではキャッシュ分割単位1302の同じグループのものを選択する。もしなければ、違うグループのものから選択してよい。次に記憶装置システム種別1303を用いて選択する。この例では、第1の記憶装置システムから選択する（1303）。但し、第1の記憶装置システム内に空きLUがなければ、第2の記憶装置システムから選択する。尚、項目や順位はここであげた例に限らない。ユーザが項目を入力してもよい。また、ユーザが条件指定する方法は、図9のステップ903で示す正LUの特性が複数ある場合の、別の特性選択方法としても使用できる。

【0061】

図11は、基準テーブル1400の一例を示す図である。本テーブルはメモリ22に格納され、ユーザが副LUの選択基準を入力した場合の、入力した項目による判定の仕方の情報を格納している。具体的には、ユーザが入力した基準1401とその場合の副LU選択アルゴリズム1402がある。例えば、ユーザが「信頼性重視」を入力すれば、判定アルゴリズム1402を参照し、FC記憶装置の中から選択する。もし、FC記憶装置に空きLUがなければ、他の記憶装置から選択する。ユーザが「バックアップ」「コピー数7」を入力すれば、判定アルゴリズム1402を参照し、ATA記憶装置の中から選択する。

【0062】

副LUを選択したら、ペア管理テーブル46に新規にエントリを作成し、ユーザが指定したパラメータ及び選択された副LUを登録する。

【0063】

このように、ユーザは正LUを選択して指定するだけで、記憶装置システム側

で最適な副LUを選択し、ペアを形成することが可能となる。また、上記フローでは候補となるLUを1つ選択しているが、複数の候補をユーザに提示し、ユーザに選択させる方法もある。

【0064】

上記説明では副LU候補は既にシステム内にあるLUから選択していた。もう一つの方法として、レプリケーションを要求するホストに空きLUがない場合、副LUを生成して割り当てることを行う。図12にその処理フローを示す。

【0065】

ステップ801の判定と同様（ステップ1201）、候補となる空きLUがあれば図8と同様に行う（ステップ1203）。空きLUがない場合、正LUの特性を検索し、正LUと同じ記憶装置種別のLUを生成する。例えば、正LUがFC記憶装置のLUであった場合は、FC記憶装置にパリティグループをきってLUを作成し、ホストにこのLUを割り当てる（ステップ1202）。副LUとして、作成したLUを割り当てる（ステップ1204）。割り当てたLUをLU管理テーブル45、及びペア管理テーブル46に登録する（ステップ1205）。この処理はステップ807と同様である。

【0066】

図7は、記憶装置システム70Aがプロトコル変換アダプタ600を有する場合の構成を示す図である。ここでは、ネットワーク60を介して、記憶装置システム70Aと他の記憶装置システム70Bが接続されている。ネットワーク60と記憶装置システム70Bがない構成でもよい。プロトコル変換アダプタ600は、記憶制御装置20と独立したチャネル接続部分であり、LAN（Local Area Network）、公衆回線、専用回線、ESCON（Enterprise Systems Connection）に準拠したプロトコルなどを扱う。複数のプロトコル変換アダプタ600と複数の記憶制御装置20はネットワーク63を介して接続されている。

【0067】

プロトコル変換アダプタ600は、ホスト10から入出力要求を受け取ると、コマンドを解析してプロトコル変換し、前記コマンドが要求するデータが格納されているLUが記憶制御装置20の配下に管理されているか記憶装置システム7

0 Bに管理されているかを判断し、判断した方に当該要求を送信する。どちらに管理されているかは、ネットワーク 63で接続されたプロセッサ 900内のメモリ上に格納された構成情報テーブル 48を参照することで判断する。記憶制御装置 20は要求を受け取ったら、副LU選択支援プログラム 43を起動する。

【0068】

管理サーバ 100は、ネットワーク 62を介して、第一の記憶装置システム 70Aを認識する。また、管理サーバ 100を、第一の記憶装置システム 70Aと専用線で直接接続する構成もある。

【0069】

図7の構成を本発明に適用することができる。

【0070】

尚、記憶装置は磁気媒体を使うことが多いが、光学的媒体などの他の媒体を利用する装置であってもよい。また、本発明で説明したプログラムは、CD-ROM等の記憶媒体から移してもよいし、ネットワーク経由で他の装置からダウンロードしてもよい。

【0071】

【発明の効果】

本発明によれば、レプリケーションを作成する際、コピー元の特性を考慮してコピー先を決定する記憶装置システムを提供することができる。また、ユーザの負担を軽減するため、ユーザが前記特性を意識せずにコピー先を選択できるようなレプリケーション支援機能を有する記憶装置システムを提供することができる。

【図面の簡単な説明】

【図1】 計算機システムの構成図。

【図2】 記憶装置管理テーブルの例を示す図。

【図3】 LU管理テーブルの例を示す図。

【図4】 ペア管理テーブルの例を示す図。

【図5】 キャッシュグループ情報テーブルの例を示す図。

【図6】 構成情報管理テーブルの例を示す図。

【図 7】 図 1 とは別の計算機システムの構成図。

【図 8】 副 L U を自動的に選択する処理フロー。

【図 9】 他の特性グループから副 L U を選択する処理フロー。

【図 10】 特性の優先順位情報テーブルの例を示す図。

【図 11】 ユーザが副 L U の選択基準を入力した項目による判定の仕方を記載する基準テーブルの例を示す図。

【図 12】 L U を生成して副 L U に割り当てる処理フロー。

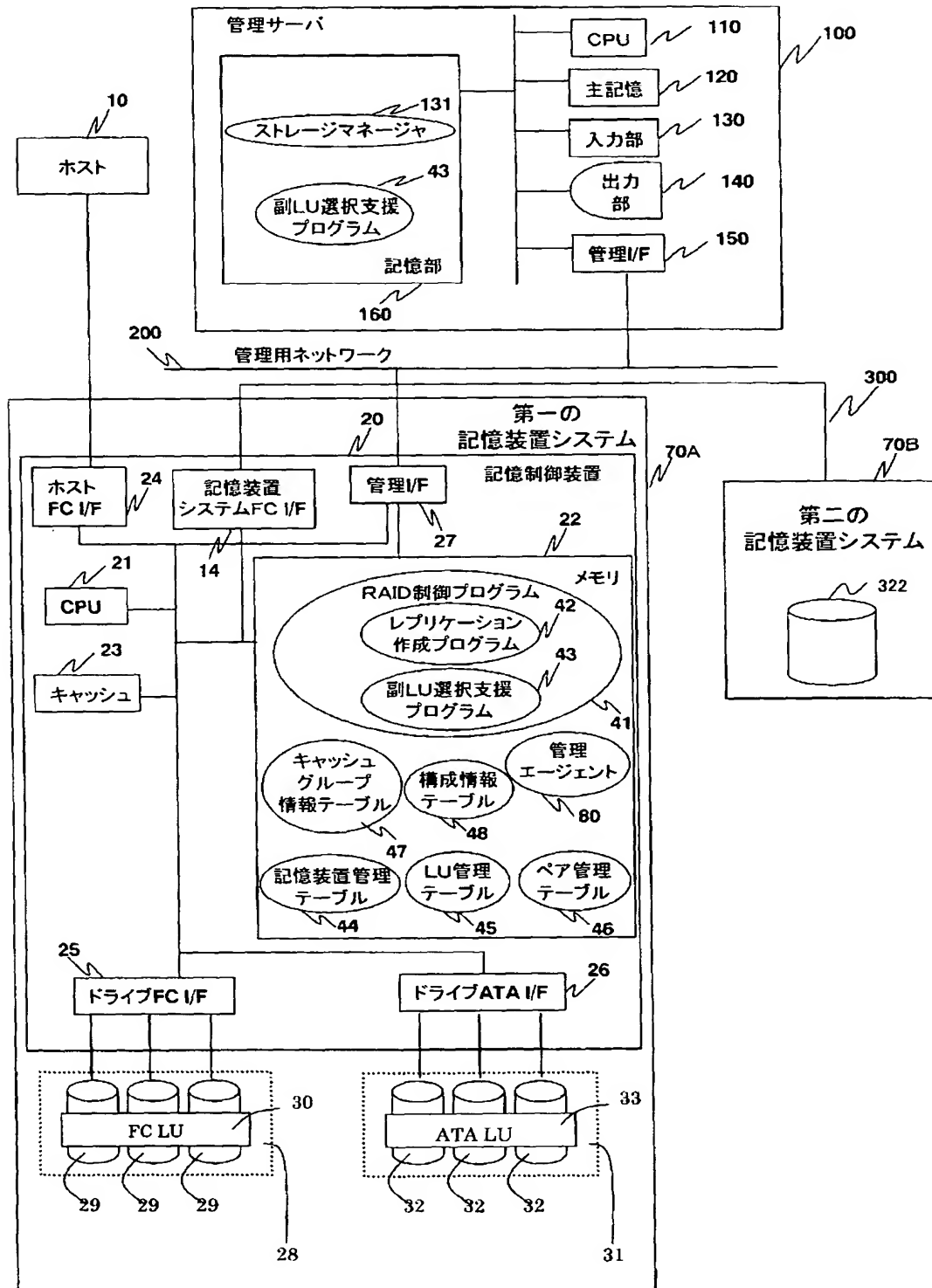
【符号の説明】

70A、70B…記憶装置システム、100…管理サーバ、10…ホスト、20…記憶制御装置、14…記憶装置システムFC I/F、27…管理 I/F、21…CPU、22…メモリ、23…キャッシュ、24…ホストFC I/F、25…ドライブFC I/F、26…ドライブATA I/F。

【書類名】 図面

【図 1】

図 1



【図 2】

図2

ドライブ 番号	ドライブ 種別	アレイ 構成	用途	起動 状況
0	FC	(1)RAID5	DB	ON
1	FC	(1)RAID5	DB	ON
2	FC	(1)RAID5	DB	ON
3	ATA	(2)RAID5	FS	ON
4	ATA	(2)RAID5	FS	ON
5	ATA	(2)RAID5	FS	ON

241 242 243 244 245

44

【図 3】

図3

LU 番号	ホスト 割当状況	LUN	容量	LU 種別	相手LU 番号
0	有	1	50GB	FC	20、158
1	無	—	50GB	FC	
2	有	1	50GB	ATA	478、1592
3	無	—	50GB	ATA	

251 252 253 254 255 256

45

【図 4】

図4

LUペア 番号	正LU情報 LU番号	副LU情報 LU番号	ペア 状態
0	100	200	Pair
1	110	210	Pair
2	120	220	Split
3	130	230	Pair
4	140	240	Split

261 262 263 264

46

【図 5】

図5

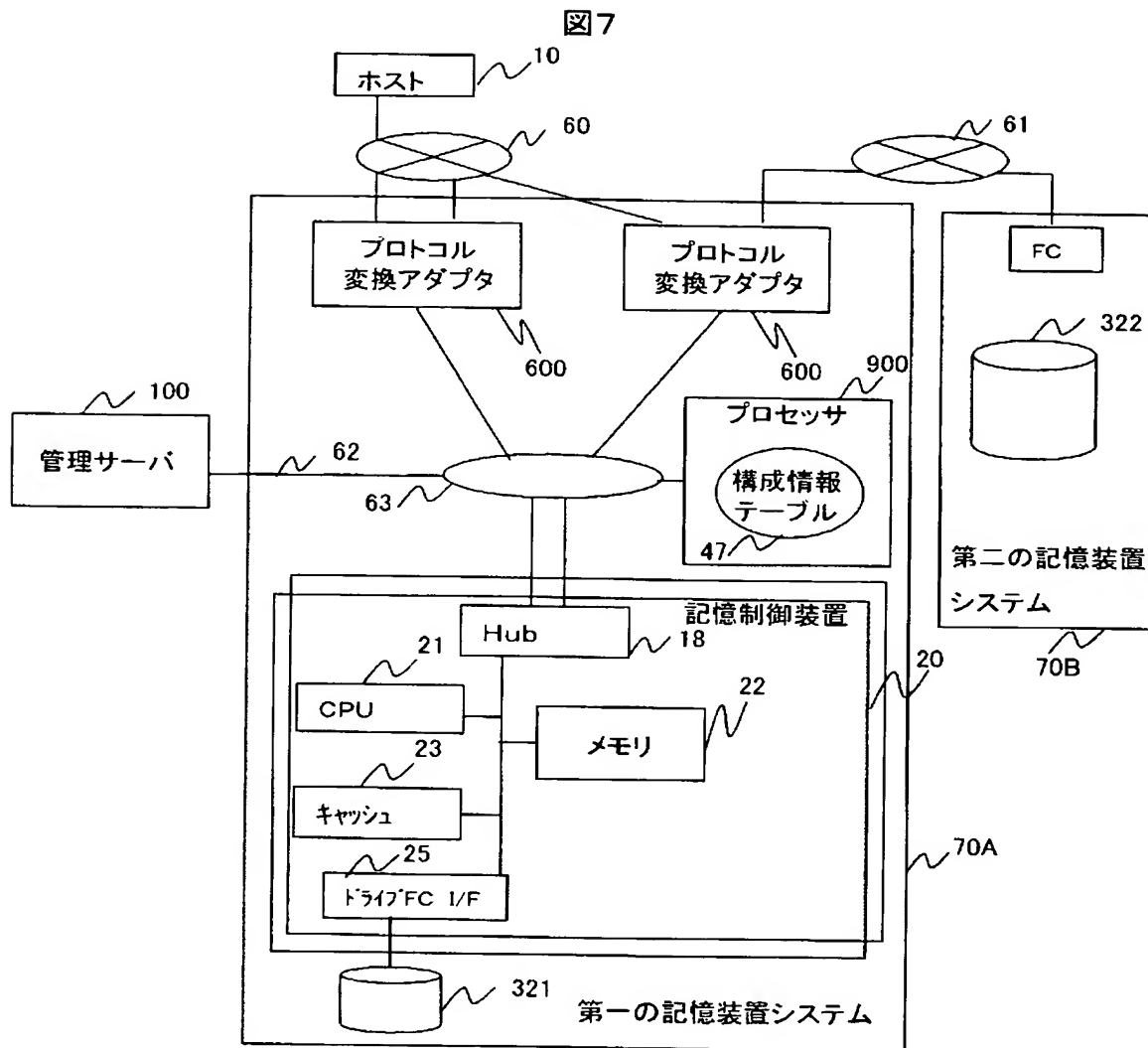
460				
キャッシュグループ ID	CGrp0	CGrp1	...	461
キャッシュ量	1024MB	2048MB	...	462
LU ID	LU 0, LU 2, LU 3, ...	LU 1	...	463
キャッシュグループ情報テーブル				

【図 6】

図6

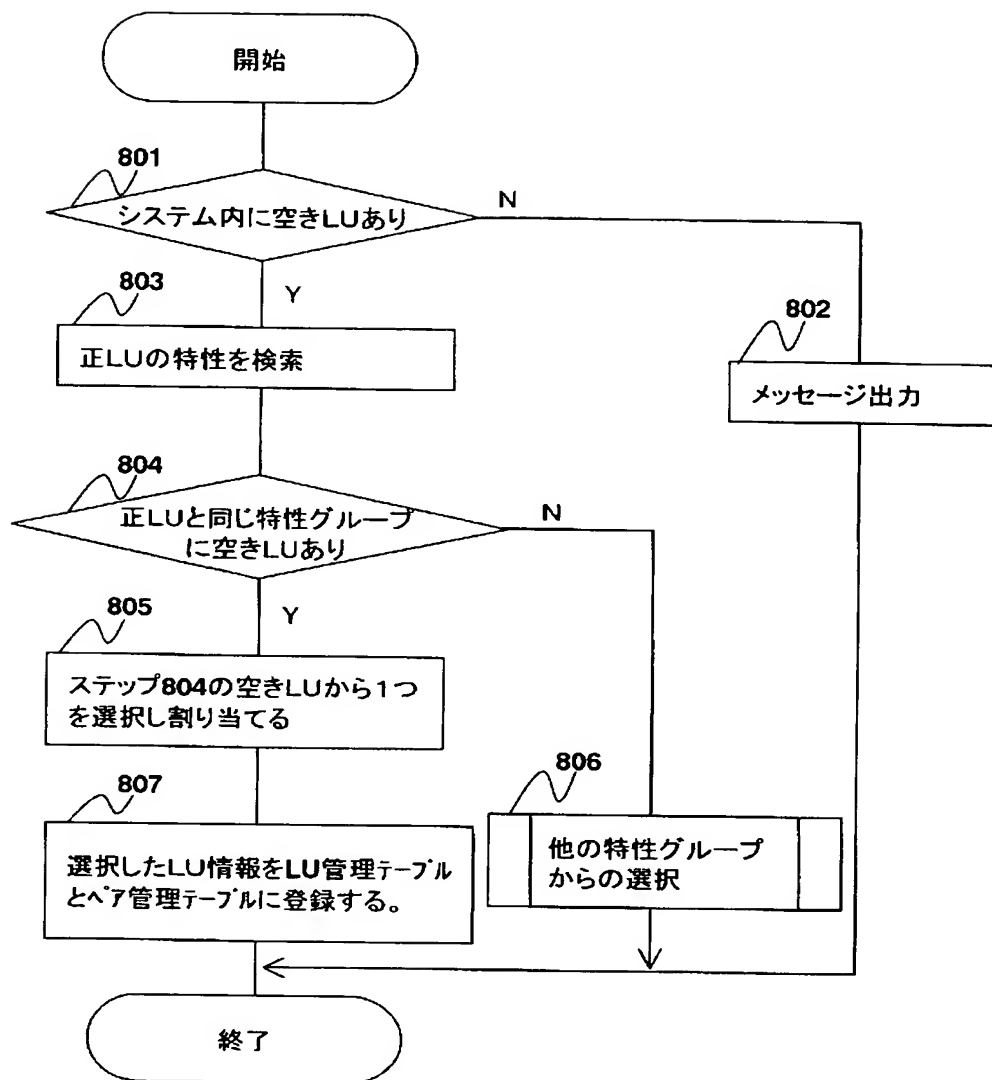
ポート ID	WWN	LUN	容量	マッピング LUN	48
010001	ad00bfz0afd	5	5120KB	3B:20	
481	482	483	484	485	

【図 7】



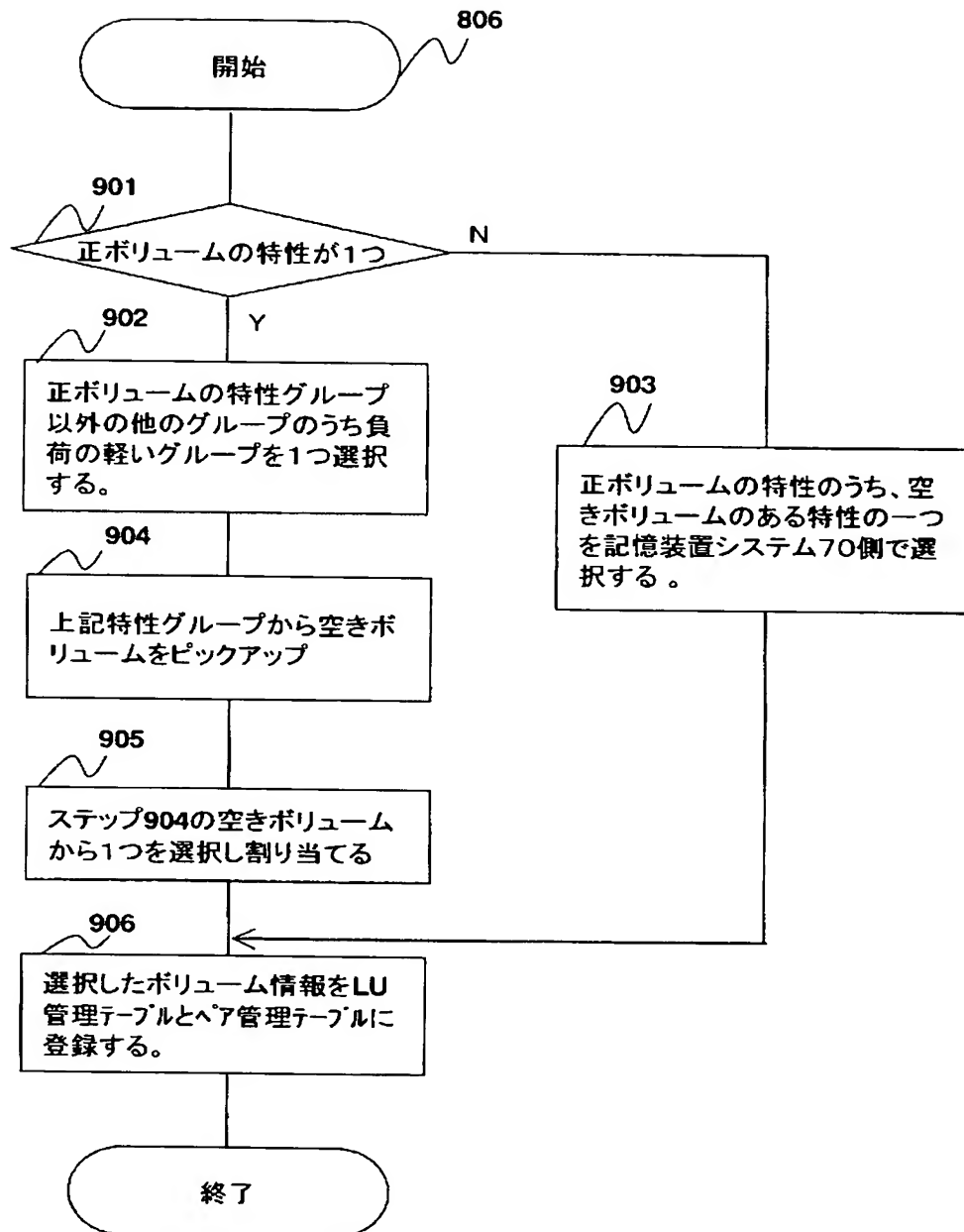
【図 8】

図8



【図 9】

図 9



【図 10】

図 10

1300	
優先順位	項 目
1	ハードディスク種別
2	キャッシュ分割単位
3	記憶装置システム種別
	第一の記憶装置システム

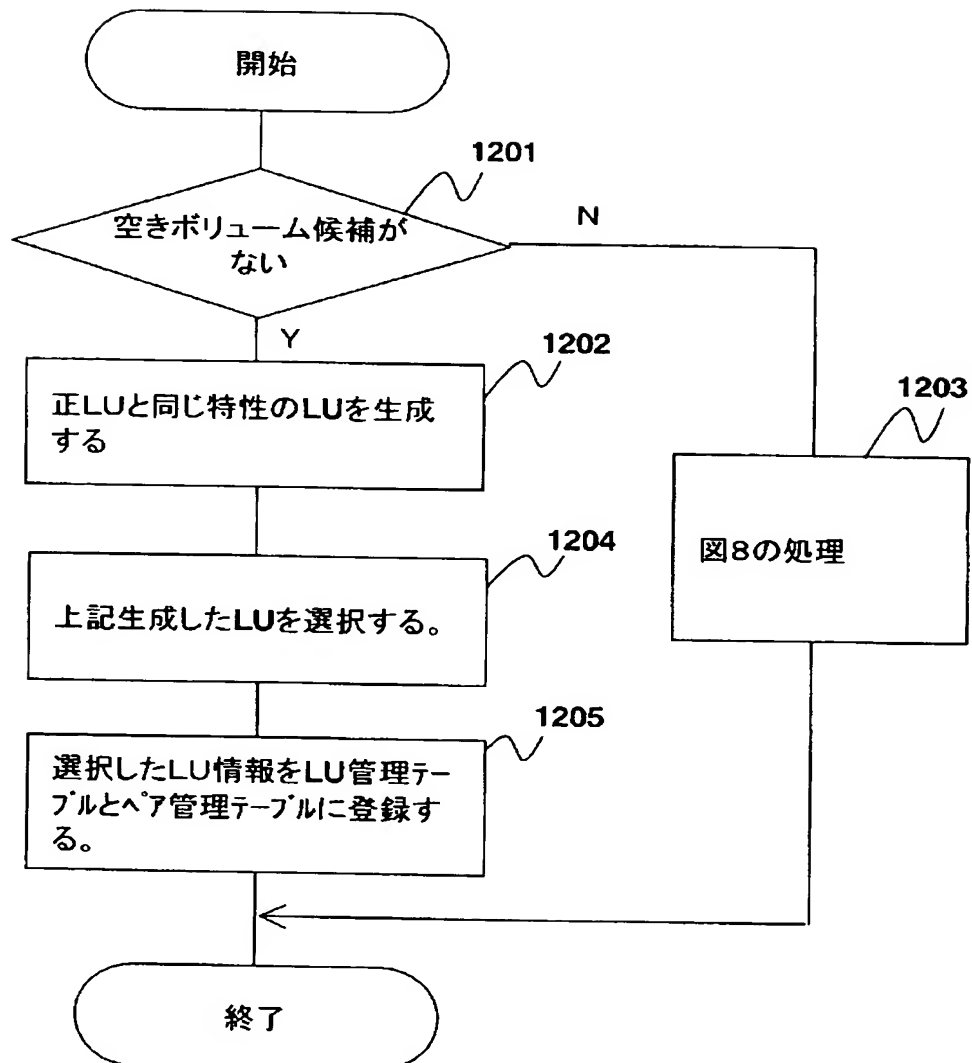
【図 11】

図 11

1400	
基準項目	判定のアルゴリズム
信頼性重視	FC記憶装置を選択
バックアップ	バックアップ且つコピー数が6以上であればATA記憶装置を選択 コピー数が5以下であれば、コピー元の記憶装置と同じものを選択
用途がDB	FC記憶装置を選択
性能重視	キャッシュ分割単位が同じ且つFC記憶装置を選択
価格重視	ATA記憶装置を選択

【図 12】

図 12



【書類名】 要約書

【要約】

【課題】 1 台の記憶装置システム内に、記憶装置種別、キャッシュ分割単位などが同じ特性グループに混在する場合、これらの特性を意識したレプリケーションボリュームを作成することは考慮されていなかった。

【解決手段】 各ボリュームの特性を記載したテーブルをシステム側で持ち、管理者は特性を意識しなくてもレプリケーションボリュームの候補の提示又は割当が行うことができる。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願 2 0 0 3 - 1 7 8 8 7 9
受付番号	5 0 3 0 1 0 4 6 9 4 2
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 5 年 6 月 2 5 日

< 認定情報・付加情報 >

【提出日】 平成15年 6月24日

次頁無

出証特 2 0 0 3 - 3 0 6 1 6 4 8

特願 2 0 0 3 - 1 7 8 8 7 9

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1 . 変更年月日
[変更理由]
住 所
氏 名

1 9 9 0 年 8 月 3 1 日
新規登録
東京都千代田区神田駿河台 4 丁目 6 番地
株式会社日立製作所